STRengthening Analytical Thinking for Observational Studies (STRATOS): Guidance for analysts with limited statistical knowledge

Georg Heinze1, Anne-Laure Boulesteix2, Daniela Dunkler1, Mitchell H. Gail3, Katherine J. Lee4, Ben van Calster5, Michael Wallace6, Willi Sauerbrei7

1 Medical University of Vienna, Center for Medical Data Science, Section for Clinical Biometrics; Vienna, Austria

2 LMU Munich, Institute for Medical Information Processing, Biometry and Epidemiology; Munich, Germany

3 National Cancer Institute, Division of Cancer Epidemiology and Genetics, Biostatistics Branch; Rockville, Maryland, USA

4 University of Melbourne, Murdoch Children's Research Institute and Department of Paediatrics, Clinical Epidemiology and Biostatistics Unit; Melbourne, Australia

5 KU Leuven, Department of Development and Regeneration; Leuven, Belgium

6 University of Waterloo, Department of Statistics and Actuarial Science; Waterloo, Canada

7 Medical Center - University of Freiburg, Institute of Medical Biometry and Statistics; Freiburg, Germany

Introducing the STRATOS initiative, Sauerbrei et al (2014) stressed that many methodological developments are not implemented in practice. Lack of guidance or recommendations on practical issues could be one important reason. Accordingly, Sauerbrei et al stated: 'an efficient way to help researchers to keep up with recent methodological developments is to develop guidance documents that are spread to the research community at large', and they defined three groups of analysts with different levels of statistical knowledge that guidance could be targeted towards. Level 1 are researchers with only basic statistical knowledge and limited experience in using statistical methodology and software. Level 2 are experienced statisticians/analysts, and level 3 are statistical experts in a specific area. Many, if not most, analyses are conducted by level 1 researchers. Recently in the Biometric Bulletin, we mentioned the aims of guidance and educational material being developed by STRATOS (BB, 2022, 39(3)), and here we briefly summarize some specific STRATOS projects tailored for level 1 analysts.

Papers explaining statistical methods in medical journals

Some medical journals publish a series of articles about statistical methods, and these articles are a primary source of guidance for level 1 analysts. A review project by topic group (TG) 2 assessed 57 papers published in 23 different series for their contents. The review identified relevant gaps in the educational material for the selection of variables and functional forms for multivariable models (Wallisch et al, 2022). Other topic groups plan to assess these articles from their perspectives, which will stimulate STRATOS' activities in contributing to such a series.

Overview and guidance papers

Recently TG1 published a framework for the treatment and reporting of missing data in observational studies (Lee et al, 2021). The paper included Stata code as well as sample text for reporting and has been presented at numerous conferences.

TG3's 'Ten simple rules for Initial Data Analysis (IDA)' offers a brief overview of the aspects of IDA and the benefits of integrating IDA into the research practice. That article addresses a broad audience (Baillie et al, 2022).

TG4's article 'Analysis in an Imperfect World' (Wallace, 2020) provides an introduction to the topic of measurement error for a general audience and was highlighted in the annual anthology Best Writing on Mathematics (Pitici, 2021).

TG5 outlined principles for designing an appropriate observational study for measuring

an association between an exposure and disease incidence (Gail et al, 2019), where the selection of an appropriate study design must often balance scientific study aims and practical constraints. The paper compared various cohort and case-control designs and discussed how certain design features can reduce threats to study validity.

TG6 explained important but often neglected, issues in developing or validating prediction models to a medical audience. For example, Wynants et al (2019) discussed common myths about choosing and interpreting risk thresholds to recommend patients for treatment, and van Calster et al (2019) discussed the importance of calibration assessment.

STRATOS also includes Panels that work on special topics. The Simulation Panel introduced basic principles for statistical simulation studies (Boulesteix et al., 2020). That paper targeted level 1 data analysts and researchers with little or no practical experience with simulations, who either (i) rely on previously published simulation studies to choose their statistical methods or (ii) wish to perform their simulation.

Videos

Led by Rolf Groenwold and Michael Wallace, STRATOS has started to produce a series of short videos about aspects of statistical analysis, already covering 'Categorisation of continuous predictors', and 'Modelling continuous predictors`, and a three-part series describing what measurement error is, why it matters for statistical analysis, and some general approaches for addressing the problems it can cause. More videos, also from other TGs, are planned.

Interactive Shiny-Apps

The possibility to develop and publish interactive apps based on statistical software makes R shiny a preferred technology for complementing guidance documents with interactive elements. A shiny app that interactively explains splines and fractional polynomials modeling, 'Bend your (sp)line' (https://clinicalbio-metrics.shinyapps.io/Bendyourspline/), was developed under the supervision of Georg Heinze, and is intended for use in teaching and consulting situations. Similarly, Veronika Deffner and colleagues from TG4 developed an app to explore the impact of measurement error and misclassification on linear regression models (https://mem-explorer.shinyapps.io/MEMExplorer-v5/). Experiences from these two projects will help with developing shiny apps for other topics. For example, TG2 is developing a shiny app on variable selection.

Plans for the future

STRATOS will soon put all educational material, including papers, videos, and shiny apps on a new landing page for interested level-1 analysts. Here, we will also provide links (with brief comments) to useful papers and other resources that were developed outside of STRATOS. Please support us by sending material YOU consider as being suitable for level 1 analysts and which is closely related to issues from one of the nine topic groups TG1-TG9 (https://www.stratos-initiative.org/). STRATOS members will assess the suitability before linking the material. Check the STRATOS website regularly for updates.

Furthermore, STRATOS intends to continue and increase its presence at non-statistical conferences. For example, TG3 has recently co-organized a workshop on data quality and IDA with 130 interdisciplinary participants that took place in Berlin, Germany, in November 2022 (BB, 2022, 39(4)).

References

Baillie, M., le Cessie, S., Schmidt, C. O., Lusa, L., & Huebner, M. (2022). Ten simple rules for initial data analysis. PLOS Computational Biology (Vol. 18, Issue 2, p. e1009819). https://doi.org/10.1371/journal.pcbi.1009819

Boulesteix, A.-L., Groenwold, R. H., Abrahamowicz, M., Binder, H., Briel, M., Hornung, R., Morris, T. P., Rahnenführer, J., & Sauerbrei, W. (2020). Introduction to statistical simulations in health research. BMJ Open (Vol. 10, Issue 12, p. e039921). https://doi.org/10.1136/bmjopen-2020-039921

Gail, M. H., Altman, D. G., Cadarette, S. M., Collins, G., Evans, S. J., Sekula, P., Williamson, E., & Woodward, M. (2019). Design choices for observational studies of the effect of exposure on disease incidence. BMJ Open (Vol. 9, Issue 12, p. e031031). https://doi.org/10.1136/bmjopen-2019-031031

Lee, K. J., Tilling, K. M., Cornish, R. P., Little, R. J. A., Bell, M. L., Goetghebeur, E., Hogan, J. W., & Carpenter, J. R. (2021). Framework for the treatment and reporting of missing data in observational studies: The Treatment And Reporting of Missing data in Observational Studies framework. Journal of Clinical Epidemiology (Vol. 134, pp. 79–88). https://doi.org/10.1016/j.jclinepi.2021.01.008

Sauerbrei, W., Abrahamowicz, M., Altman, D. G., Cessie, S., & Carpenter, J. (2014). STRengthening Analytical Thinking for

Observational Studies: the STRATOS initiative. Statistics in Medicine (Vol. 33, Issue 30, pp. 5413–5432). https://doi.org/10.1002/sim.6265

Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: the Achilles heel of predictive analytics. BMC Medicine (Vol. 17, Issue 1). https://doi.org/10.1186/s12916-019-1466-7

Wallace, M. (2020). Analysis in an imperfect world. Significance (Vol. 17, Issue 1, pp. 14–19). https://doi.org/10.1111/j.1740-9713.2020.01353.x

Wallisch, C., Bach, P., Hafermann, L., Klein, N., Sauerbrei, W., Steyerberg, E. W., Heinze, G., & Rauch, G. (2022). Review of guidance papers on regression modeling in statistical series of medical journals. PLOS ONE (Vol. 17, Issue 1, p. e0262918). https://doi.org/10.1371/journal.pone.0262918

Wynants, L., van Smeden, M., McLernon, D. J., Timmerman, D., Steyerberg, E. W., & Van Calster, B. (2019). Three myths about risk thresholds for prediction models. BMC Medicine (Vol. 17, Issue 1). https://doi.org/10.1186/s12916-019-1425-3

Region News

Australasian Region (AR)

Annual general meeting

We had 21 members attend our Annual general meeting on 23 November 2022. We thanked our outgoing president Dr. Vanessa Cave for the leadership she contributed to the society over the past two years. We also welcomed our new president, James Curran. James is a Professor of Statistics and Head of the Department at the University of Auckland. James has a strong interest in statistical problems in Forensic Science, Statistical Computing, and Data Science. He is a past President of the New Zealand Statistical Association and the New Zealand Forensic Science Society. He also currently serves as the president of the Australian and New Zealand Forensic Science Society.

Professor James Curran, newly elected president for the Australasian Region of the International Biometric Society.

Congratulations to Ken Dodds

Congratulations to Ken Dodds, AgResearch, and IBS-AR member, on having been awarded the Jones Medal, for lifetime achievement in pure or applied mathematics or statistics, by the Royal Society Te Apārangi (Royal Society of New Zealand) for his work developing and applying statistical methods for genetic data analysis that enable the use of low-cost genotyping in primary industries and ecology.

IBS-AR member Ken Dodds received the Jones Medal from the Royal Society Te Apārangi.

IBS-AR conference

For your diary: the joint IBS-AR/SEEM Regional Conference, November 27 to December 1, 2023, at the Copthorne Hotel and Resort, Bay of Islands, Waitangi, New Zealand. It will be a joint conference with the Statistics in Ecology and Environmental Monitoring (SEEM) meeting. See https://biometricsboi.nz/ for details.

Garth Tarr
Biometric Bulletin Correspondent

Brazilian Region (RBras)

RBras 2022 annual Conference

The 66th Annual Meeting of the Brazilian Region of the International Society of Biometrics (RBras) was held between 16th and 18th November 2022, in Florianópolis, Santa Catarina, Brazil. Professors, Assistant professors, and Senior Lecturers from the Federal University of Santa Catarina and the Federal University of Lavras served on the event's Local Organising Committee.

The scientific program included Brazilian and international keynote speakers, short courses, and invited sessions, as well as oral and poster presentations.

The Annual Meeting of RBras is one of the most important Brazilian scientific congresses in statistics and biometrics, and in 2022 it was attended by 190 people from Brazil and abroad, including students, lecturers, professors, researchers, and professionals in statistics, biometrics, applied statistics, and data science.